

SEQUENCE LISTING

<110> INVITROGEN CORPORATION
DALBY, Brian
BENNETT, Robert

<120> DELIVERY OF FUNCTIONAL PROTEIN SEQUENCES
BY TRANSLOCATING POLYPEPTIDES

<130> INVIT1280-1

<150> PCT/US00/08571
<151> 2000-03-31

<150> 60/127,467
<151> 1999-03-31

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 6404
<212> DNA
<213> Artificial Sequence

<220>
<223> vector pVP22/Myc-His

<400> 1
gacggatcgg gagatctccc gatcccstat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgttt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggtagt gcgccccggc ctgcttcggc atgtacgggc cagatatacg cgttgcacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cggtacataa cttacggtaa atggcccgcc tggctgaccg cccaaacgacc 360
cccgccccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacccctt tgggacttcc ctacttggca gtacatctac gtattagtca 600
tcgcttattac catggtgatg cgggtttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tggaggtttg tttggcacc 720
aaaatcaacg ggacttcca aatgtcgtt acaactccgc cccattgacg caaatggcg 780
gtaggcgtgt acgggtggag gtctatataa gcagagctct ctggcttaact agagaaccca 840
ctgcttactg gcttatacgaa attaatacga ctcactatag ggagacccaa gctggcttagt 900
taagcttatt atgacctctc gccgctccgt gaagtcgggt cccggggagg ttccgcgcga 960
tgagtacgag gatctgtact acaccccttc ttcaaggatg gcgagttccg atagtccgc 1020
tgacacccctcc cggccgtggcg ccctacagac acgctcgcgc cagagggcg agtccgttt 1080
cggtccgtac gacgagtcgg attatgcctt ctacggggc tcgtcttccg aagacgacga 1140
acacccggag gtccccccggc cgcggcgtcc cgtttccggg gcggtttgt cccggccggg 1200
gcctgcgcgg ggcgcctccgc caccggctgg gtccggaggg gccggacgc caccaccac 1260
cgccccccgg gccccccggaa cccagcgggt ggcgtctaag gccccccggg cccggcggc 1320
ggagaccacc cgcggcagga aatcgccca gccagaatcc gccgcactcc cagacgcccc 1380
cgcgctcgacg ggcggcaaccc gatccaagac accccgcgcag gggctggcca gaaagctgca 1440
ctttagcacc gcccccccaa accccggacgc gccatggacc ccccggtgg cccgctttaa 1500
caagcgcgttc ttctgcgcgg cggtcggggcg cctggcggcc atgcgtccc ggtggcggc 1560
tgtccagctc tgggacatgt cgcgtccggc cacagacgaa gacccaacg aactccttgg 1620
catcaccacc atccgcgtga cggtctgcga gggcaaaaac ctgcttcagc ggcggcaacga 1680
gttggtaat ccagacgtgg tgcaggacgt cgacgcggcc acggcgactc gagggcgttc 1740

tgcggcgtcg	cgccccaccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgccccag	1800
acggcccgtc	gagggtaccg	agctcgatc	cactagtcca	gtgtgggtga	attctgcaga	1860
tatccagcac	agtggcgcc	gctcgagtct	agagggccc	cggttcaac	aaaaactcat	1920
ctcagaagag	gatctgaata	tgcataccgg	tcatcatcac	catcaccatt	gagttaaac	1980
ccgctgatca	gcctcgactg	tgccttctag	ttgcccagcca	tctgtgttt	gccctccccc	2040
cgtgccttcc	ttgaccctgg	aagggtgccac	tcccactgtc	ctttcctaata	aaaatgagga	2100
aattgcatacg	cattgtctga	gttaggtgtca	ttctattctg	gggggtgggg	tggggcagga	2160
cagcaagggg	gaggattggg	aagacaatacg	caggcatgt	ggggatgcgg	tgggctctat	2220
ggttctgag	gcccggaaagaa	ccagctgggg	ctctaggggg	tatccccacg	cggccctgttag	2280
cggcgcattta	agcgcggcg	gtgtgggtgt	tacgcgcage	gtgaccgcta	cacttgcac	2340
cggccctagcg	cccgccctt	tcgctttttt	ccctccctt	ctcgccacgt	tcgcccgtt	2400
tccccgtca	gctctaaatc	ggggcatccc	tttagggttc	cgatttagtgc	cttacggca	2460
cctcgacccc	aaaaaaacttg	attagggtga	tggttcacgt	agtggccat	cggccctgtata	2520
gacggttttt	cgcccctttga	cggtggagtc	cacgttcttt	aatagtggac	tcttggtcca	2580
aactggaaaca	acactcaacc	ctatctcggt	ctattttttt	gatttataaag	ggattttggg	2640
gatttccggcc	tattgtttaa	aaaatgagct	gatttaacaa	aaatttaaacg	cgaattaatt	2700
ctgtggaatg	tgtgtcagtt	agggtgtgg	aagtccccag	gctcccccagg	caggcagaag	2760
tatgcaaaagc	atgcatactca	attagtca	aaccagggtgt	ggaaagtccc	caggctccccc	2820
agcaggcaga	agtatgcaaa	gcatgcattc	caattagtca	gcaaccatag	tcccggccct	2880
aactccgcccc	atcccccccc	taactccgccc	cagtccgccc	cattctccgc	ccatggctg	2940
actaattttt	tttattttatg	cagaggccga	ggccgcctct	gcctctgagc	tattccagaa	3000
gtagtgagga	ggctttttt	gaggccttagg	ctttgcaaa	aagctcccg	gagcttgat	3060
atccattttc	ggatctgatc	aagagacagg	atgaggatcg	tttgcgtat	ttgaacaaga	3120
tggattgcac	gcaggttctc	cgcccgctt	ggtggagagg	ctattcggt	atgactggc	3180
acaacagaca	atcggtctgt	ctgatggcc	cgtgttccgg	ctgtcagcgc	aggggcgccc	3240
ggttctttt	gtcaagaccg	acctgtccgg	tgccctgaat	gaactgcagg	acgaggcagc	3300
gcccgtatcg	tggctggcca	cgacggggct	tccttgcgc	gctgtgtcg	acggtgtcac	3360
tgaagcggga	agggactggc	tgctattggg	cgaagtgcgg	gggcaggatc	tcctgtcatc	3420
tcaccttgct	cctgcccaga	aagtatccat	catggctgtat	gcaatgcggc	ggctgcatac	3480
gcttgcattcg	gctacactgccc	cattcgacca	ccaaagcgaaa	catcgatcg	agcgagcacg	3540
tactcgatcg	gaagccggc	ttgtcgatca	ggatgatctg	gacgaagagc	atcaggggct	3600
cgcgcctagcc	gaactgttc	ccaggctcaa	ggcgcgcatg	cccgacggcg	aggatctcg	3660
cgtgacccat	ggcgatgcct	gcttgcggaa	tatcatggtg	aaaaatggcc	gctttctgg	3720
attcatcgac	tgtggccggc	ttgggtgtggc	ggaccgctat	caggacatag	cggtggctac	3780
ccgtgatatt	gctgaagagc	ttggccggca	atgggctgac	cgcttcctcg	tgctttacgg	3840
tatcgccgct	cccgatttc	agcgcatacg	cttctatcg	cttcttgacg	atttcttctg	3900
agcgggactc	tggggttcgc	aaaatgaccc	accaagcgac	gccccacctg	ccatcacgag	3960
atttcgattc	caccggcc	ttctatgaaa	ggttgggctt	cggaatcg	ttccggacg	4020
ccggctggat	gatccctcc	cgccggggatc	tcatgtgg	gttcttcg	caccccaact	4080
tgtttattgc	agcttataat	ggttacaat	aaagcaatag	catcacaat	ttcacaatata	4140
aagcattttt	ttcactgc	tctagttgt	gtttgcctaa	actcatcaat	gtatcttac	4200
atgtctgtat	accgtcgacc	tctagctaga	gcttggcgta	atcatggtca	tagctgtttc	4260
ctgtgtaaa	ttgttatcc	ctcacaattc	cacacaacat	acgagccg	agcataaaat	4320
gtaaagcctg	gggtgcctaa	tgagttagct	aactcacatt	aattgcgttgc	cgctca	4380
ccgcttccca	gtcggaaac	ctgtcg	agctgcatta	atgaatcg	caacgcgcgg	4440
ggagaggcgg	tttgcgtatt	gggcgtctt	ccgcttc	gctca	tcgctgcgt	4500
cgtcgatcg	gctcgccg	ggggatctag	ctcactcaaa	ggcggtata	cggttatcca	4560
cagaatcagg	ggataacgc	ggaaagaaca	tgtgagcaaa	aggccagca	aaggccagga	4620
accgtaaaaaa	ggccgcgtt	ctggcg	tccataggct	ccggccccct	gacgagcatc	4680
acaaaaatcg	acgctcaagt	cagagggtgg	gaaacccgac	aggactataa	agataccagg	4740
cgttccccc	ttggaa	ctcgatcg	ctccgttcc	gaccctgc	cttaccggat	4800
acctgtccgc	ctttccct	tcggaaagcg	ttggcg	tcaatgc	cgctgttaggt	4860
atctcgttgc	gggtgttagt	ttcgatcca	agctggctg	tgtgcac	ccccccgttc	4920
agcccgaccg	ctgcgcctt	tccgttaact	atcgcttgc	gtccaacccg	gtaagacacg	4980
acttatacgcc	actggcagca	gccactggta	acaggattag	cagagcg	tatgttaggc	5040
gtgctacaga	tttgcgtt	ttggcg	actacggct	cactagaagg	acagtatttg	5100
gtatctgcgc	tctgtgt	ccagttact	tcggaaaaag	agttggtagc	tcttgcgtt	5160
gcaaacaac	caccgcgtt	agcgggtgg	ttttgtttt	caagcagc	attacgcgc	5220
aaaaaaaagg	atctcaagaa	gatccctt	tctttctac	ggggctgtac	gctcagtgg	5280
acaaaaactc	acgttaaggg	attttggtca	tgagattatc	aaaaggatc	ttcacctt	5340
tccttttaaa	ttaaaaatga	agttttaat	caatctaaag	tatataatgag	taaacttgg	5400

ctgacagtta	ccaatgctta	atcagtggagg	cacctatctc	agcgatctgt	ctatttcgtt	5460
catccatagt	tgccctgactc	cccgtcggt	agataactac	gatacgggag	ggcttaccat	5520
ctggccccag	tgctgcaatg	ataccgcgag	acccacgctc	accggctcca	gatttatca	5580
caataaacc	gccagccgga	agggccgagc	gcagaagtgg	tcctgcaact	ttatccgcct	5640
ccatccagtc	tattaattgt	tgccgggaag	ctagagtaa	tagttcgcca	gttaatagtt	5700
tgcgcaacgt	tgttggcatt	gtacagggca	tcgtgggtgc	acgctcg	tttggatgg	5760
cttcattcag	ctcccggtcc	caacgatcaa	ggcgagttac	atgatcccc	atgttgtca	5820
aaaaagcggt	tagctccctc	ggtcctccga	tcgtgtcag	aagtaagtt	gccgcagtgt	5880
tatcaactcat	ggttatggca	gcaactgcata	atttcattac	tgtcatgcca	tccgtaaagat	5940
gctttctgt	gactggtgag	tactcaacca	agtcattctg	agaatagtgt	atgcggcgac	6000
cgagttgctc	ttgcccggcg	tcaatacggg	ataataccgc	gccacatagc	agaactttaa	6060
aagtgcgtat	cattggaaaa	cgttcttcgg	ggcgaaaaact	ctcaaggatc	ttaccgctgt	6120
ttagatccag	ttcgatgtaa	cccactcg	cacccaactg	atcttcagca	tctttactt	6180
tcaccagcgt	ttctgggtga	gcaaaaaacag	gaaggcaaaa	tgcccaaaaa	aagggaataa	6240
ggcgacacg	gaaatgttga	atactcatac	tcttccttt	tcaatattat	tgaagcattt	6300
atcagggtta	ttgtctcatg	agcggataca	tatttgaatg	tatttagaaa	aataaaacaaa	6360
taggggttcc	gogcacattt	ccccgaaaaag	tgccacactga	cgtc		6404

<210> 2
<211> 6420
<212> DNA
<213> Artificial Sequence

<220>
<223> vector pVP22/Myc-His-TOPO

<400> 2						
gacggatcgg	gagatctccc	gatcccstat	ggtcgactct	cagtaatac	tgctctgat	60
ccgcatatgtt	aagccagtat	ctgctccctg	cttgtgttt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgc	aagaatctgc	180
ttagggtag	gcgtttcg	ctgcttcg	atgtacggc	cagatatac	cggtgacatt	240
gattattgac	tagtattaa	tagtaatcaa	ttacgggtc	attagttcat	agccatata	300
tggagttccg	cgttacataa	cttacgtaa	atggccgc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacgt	aaactgccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgcccagta	catgacctta	tgggactt	ctacttggca	gtacatctac	gtattagtca	600
tcgcttattac	catggtgatg	cggtttggc	agtacatcaa	tggcgtgga	tagcggttt	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtca	tggagttt	tttggcacc	720
aaaatcaacg	ggacttcca	aatgtcgt	acaactccgc	cccatggacg	caaatggcgc	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctc	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcaactatag	ggagacccaa	gctggctagt	900
taagcttatt	atgaccttc	gccgtccgt	gaagtgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acacccgtc	ttcaggatg	gcgagtccc	atagtccgc	1020
tgacacctcc	cggcgtggc	ccctacagac	acgctcg	cagagggcg	aggccgtt	1080
cgtccagta	gacgagtcgg	attatgcct	ctacggggc	tcgtctcc	aagacgacga	1140
acacccggag	gtccccgg	cgcggcg	cgttccgg	gcgtttt	ccggccgg	1200
gcctgcgcg	gcccctccgc	cacccgctgg	gtccggaggg	gcggac	cacccaccac	1260
cggcccccgg	gccccccgaa	cccagcggt	ggcactaag	gcccccg	ccccggcgg	1320
ggagaccacc	cgcggcagga	aatcgccca	gccagaatcc	gcgcactcc	cagacgccc	1380
cgcgtcgacg	gcccacacc	gatccaagac	acccgcgc	ggcgtggc	gaaagctgca	1440
ctttagcacc	gccccccaa	accccgacgc	gccatggacc	ccccgggt	ccggcttta	1500
caagcgcgtc	ttctgcgccc	cggtcggggc	cctggcggcc	atgcatccc	gatggcgg	1560
ggtccagctc	tgggacatgt	cgcg	cacagacaa	gaccta	aactccttgg	1620
catcaccacc	atccgcgt	cggtctgc	ggcaaaaac	ctgcttc	cgccaaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcgg	acggcactc	gagggcg	1740
tgccggcgtc	cgccccacc	agcgcac	agccccagcc	cgctccg	ctcgccccag	1800
acggcccg	gagggtacc	agctcg	caactgtca	gtgtgg	attgccc	1860
aggcaattc	tgcagat	cagcacag	gcggccg	gagtctag	ggcccg	1920
tca	actcatctca	gaagaggatc	tgaatatgc	taccgg	catcaccatc	1980
accattgagt	ttaaacc	cgactgtgc	ttctagtt	cagccat	ctg	2040

ttgtttgccc	ctccccgtg	ctttccttga	cccttggaaagg	tgccactccc	actgtccttt	2100
cctaataaaa	tgagggaaatt	gcategcatt	gtctgagtag	gtgtcattct	attctggggg	2160
gtgggggtggg	gcaggacagc	aaggggggagg	attgggaaga	caatagcagg	catgctgggg	2220
atgcgggtgg	ctctatggct	tctgaggcgg	aaagaaccag	ctggggctct	agggggatc	2280
cccacgcgc	ctgttagcggc	gcattaagcg	cggcggtgt	ggtggttacg	cgagcgtga	2340
ccgctacact	tgccagcggc	ctagcggccg	ctcccttcgc	tttcttccct	tcctttctcg	2400
ccacgttcgc	cggcttcccc	cgtcaagctc	taaatcgggg	catccctta	gggttccgat	2460
ttagtgtctt	acggcacctc	gaccccaaaa	aacttgatta	gggtgatggt	tcacgtatg	2520
ggccatcgcc	ctgatagacg	gttttgcggc	ctttgacgtt	ggagtccacg	ttctttaata	2580
gtggactctt	gttccaaact	ggaacaacac	tcaaccctat	ctcggtctat	tcttttatt	2640
tataagggtt	tttggggatt	tcggcctatt	ggttaaaaaaa	tgagctgatt	taacaaaaat	2700
ttaacgcgaa	ttaattctgt	ggaatgtgtg	tcagtttaggg	tgtggaaagt	ccccaggctc	2760
cccaggcagg	cagaagtatg	caaagcatgc	atctcaatta	gtcagcaacc	aggtgtggaa	2820
agtccccagg	ctccccagca	gdcagaagta	tgcaagcat	gcatctcaat	tagtcagcaa	2880
ccatagtccc	gccccctaact	ccgccccatcc	cgccccctaac	tccgcccagt	tccgcccatt	2940
ctccgcccc	tggctgacta	attttttta	tttatgcaga	ggccgaggcc	gcctctgcct	3000
ctgagctatt	ccagaagtag	tgaggaggt	tttttggagg	cctaggctt	tgaaaaaagc	3060
tcccgggagc	ttgtatatcc	attttcggat	ctgatcaaga	gacaggatga	ggatcgttc	3120
gcatgattga	acaagatgga	ttgcacgcag	gttctccgc	cgcttgggt	gagaggctat	3180
tcggctatga	ctgggcacaa	cagacaatcg	gctgctctga	tgccgcccgt	ttccggctgt	3240
cagcgcaggg	gcccgggtt	ctttttgtca	agaccgacct	gtccggtgcc	ctgaatgaac	3300
tgcaggacga	ggcagcgcgg	ctatcgtggc	tggccacgac	gggcgttcc	tgccgcagctg	3360
tgctcgacgt	tgtcaactgaa	gccccggagg	actggctgt	attggcgaa	gtccggggc	3420
aggatctcct	gtcatctcac	cttgcctctg	ccgagaaaagt	atccatcatg	gctgatgcaa	3480
tgccggcgct	gcatacgctt	gatccggcta	cctgcccatt	cgaccaccaa	gcaaaacatc	3540
gcatcgagcg	agcacgtact	cgatggaaag	ccggcttctgt	cgatcaggat	gatctggacg	3600
aagagcatca	ggggctcgcg	ccagccgaac	tgttcgccag	gctcaaggcg	cgcatgccc	3660
acggcgagga	tctcgtcgt	acccatggcg	atgcctgctt	gccgaatatc	atggtgaaa	3720
atggccgctt	ttctgatttcc	atcgactgt	gccggctggg	tgtggccgac	cgctatcagg	3780
acatagcggt	ggctaccctgt	gatattgctg	aagagcttgg	cgcgaaatgg	gctgaccgct	3840
tcctcgtgt	ttacgtatc	gccgctcccg	attcgcagcg	catcgccctt	tatcgccctt	3900
ttgacgagtt	cttctgagcg	gactctgtgg	gttcgcgaaa	tgaccgacca	agcgacgccc	3960
aacctgccc	cacgagattt	cgattccacc	gccgccttct	atgaaaggtt	gggcttcgg	4020
atcgtttcc	gggacgcgg	ctggatgatc	ctccagcgcg	gggatctcat	gctggagttc	4080
ttcgcacc	ccaaacttgtt	tattgcagct	tataatggtt	acaaataaaag	caatagcatc	4140
acaaatttca	caaataaaagc	attttttca	ctgcattcta	gttgtggttt	gtccaaactc	4200
atcaatgtat	cttacatgt	ctgtataccg	tcgacctcta	gctagagctt	ggcgtaatca	4260
tggcatagc	tgtttctgt	gtgaaattgt	tatccgctca	caattccaca	caacatacga	4320
gccggaagca	taaagtgtaa	agcctgggg	gcctaatgag	tgagctaact	cacattaatt	4380
gcttgcgt	caactgcccgc	tttccagtcg	ggaaacctgt	cgtccagct	gcattaatga	4440
atcgcccaac	gcccggggag	aggcggttt	cgtattggc	gctttccgc	ttcctcgctc	4500
actgactcgc	tgcgctcggt	cggtcggt	cggcgagcg	tatcagctca	ctcaaaggcg	4560
gtaatacggt	tatccacaga	atcagggat	aacgcaggaa	agaacatgt	agaaaaaggc	4620
cagaaaaagg	ccaggaacccg	taaaaaaggcc	gcgttgcgt	cgttttcca	taggctccgc	4680
ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtacga	ggtggcgaaa	cccgacagga	4740
ctataaagat	accaggcggt	tcccccttgg	agctccctcg	tgcgctctcc	tgttccgacc	4800
ctgcccgtt	ccggataacct	gtccgcctt	ctcccttcgg	gaagcgtggc	gctttctca	4860
tgctcacgt	gtaggtatct	cagttcggt	taggtcgttc	gctccaagct	ggctgtgt	4920
cacgaacccc	cggttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tctttagtcc	4980
aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	5040
gcgagggtat	taggcgtgc	tacagagtt	ttgaagtgg	ggcctaacta	cgctacact	5100
agaaggacag	tatttggat	ctgcgctctg	ctgaagccag	ttaccttccg	aaaaagagtt	5160
ggtagcttt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	tgttgcaag	5220
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	5280
tctgacgctc	agtggaaacga	aaactcacgt	taagggattt	tggcatgag	attatcaaaa	5340
aggatctca	cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	ctaaagtata	5400
tatgagtaaa	cttggctctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	5460
atctgtctat	ttcgttcatac	catagttgcc	tgactccccg	tcgtgtatg	aactacgata	5520
cgggagggtct	taccatctgg	ccccagtgt	gcaatgatac	cgcgagaccc	acgctcaccc	5580
gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	aagtggtcct	5640
gcaactttat	ccgcctccat	ccagtctatt	aatttttgc	ggaaagctag	agtaagtagt	5700

tcgccagtt	atagttgcg	caacgttgg	gccattgcta	caggcatcgt	ggtgtcacgc	5760
tcgtcgttt	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	5820
tcccccatgt	tgtcaaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	5880
aagttggccg	cagtgttatac	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	5940
atgccatccg	taagatgctt	ttctgtgact	ggtgagtaact	caaccaagtc	attctgagaa	6000
tagtgtatgc	ggcgaccgag	ttgtcttgc	ccggcgtaa	tacgggataa	taccgcgcga	6060
catagcagaa	ctttaaaagt	gctcatcatt	ggaaaaacgtt	cttcggggcg	aaaactctca	6120
agcatcttac	cgctgtttag	atccagttcg	atgtaacccca	ctcgtgcacc	caactgatct	6180
tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	6240
gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	tcataactt	cccttttcaa	6300
tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	6360
tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgtc	6420

2020-07-02 2020-07-02
 <210> 3
 <211> 12
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Peptide linker

 <400> 3
 Leu Ala Arg Leu Leu Ala Arg Leu Leu Ala Arg Leu
 1 5 10

<210> 4
 <211> 67
 <212> PRT
 <213> Artificial sequence

 <220>
 <223> Conserved sequence of steroid/thyroid hormone
 receptor superfamily DNA-binding domain

<221> VARIANT
 <222> (0)...(0)
 <223> Xaa is non-conserved amino acids within the
 DNA-binding domain

 <221> VARIANT
 <222> 7, 9, 11, 13, 22, 27, 58, 61, 66
 <223> amino acid residues that are almost universally
 conserved, but for which variations have been
 found in some identified hormone receptors

<400> 4
 Cys Xaa Xaa Cys Xaa Xaa Asp Xaa Ala Xaa Gly Xaa Tyr Xaa Xaa Xaa
 1 5 10 15
 Xaa Cys Xaa Xaa Cys Lys Xaa Phe Phe Xaa Arg Xaa Xaa Xaa Xaa Xaa
 20 25 30
 Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Lys
 35 40 45
 Xaa Xaa Arg Xaa Xaa Cys Xaa Xaa Cys Arg Xaa Xaa Lys Cys Xaa Xaa
 50 55 60
 Xaa Gly Met
 65

<210> 5
 <211> 34
 <212> DNA

<213> Bacteriophage P1

<400> 5

ataaacttcgt atagcataca ttatacgaag ttat

34

<210> 6

<211> 34

<212> DNA

<213> E. coli

<400> 6

ataaacttcgt atagtataca ttatacgaag ttat

34

<210> 7

<211> 34

<212> DNA

<213> E. coli

<400> 7

acaacttcgt ataatgtatg ctatacgaag ttat

34

<210> 8

<211> 34

<212> DNA

<213> Saccharomyces cerevisiae

<400> 8

gaagttccta ttctctagaa agtataggaa cttc

34

<210> 9

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> peptide linker moiety

<400> 9

Gly Gly Gly Gly Ser

1

5

<210> 10

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety; sequence can be repeated
indefinit number of times

<400> 10

Gly Gly Gly Gly Ser

1

5

<210> 11

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Peptide linker moiety

2020/07/20 10:56:50

<400> 11
Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser
1 5 10

<210> 12
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 12
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
1 5 10

<210> 13
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 13
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Ser Gly Ser Thr
1 5 10 15
Lys Gly

<210> 14
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 14
Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly
1 5 10

<210> 15
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 15
Gly Ser Thr Ser Gly Ser Gly Lys Pro Gly Ser Gly Glu Gly Ser Thr
1 5 10 15
Lys Gly

<210> 16
<211> 14
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 16
Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Glu Phe
1 5 10

<210> 17
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 17
Ser Arg Ser Ser Gly
1 5

<210> 18
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 18
Ser Gly Ser Ser Cys
1 5

<210> 19
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<400> 19
Met Gly Arg Ser Gly Gly Gly Cys Ala Gly Asn Arg Val Gly Ser Ser
1 5 10 15
Leu Ser Cys Gly Gly Leu Asn Leu Gln Ala Met
20 25

<210> 20
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker moiety

<221> VARIANT
<222> (0)...(0)
<223> Xaa is (GmS)_n, where m is
from 2 to 4 and n is from 1 to 11.

<400> 20
Ala Met Xaa Ala Met

1

5

<210> 21
<211> 16
<212> PRT
<213> Drosophila acanthoptera

<400> 21
Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys
1 5 10 15